Preventing Preschool Fadeout through Instructional Intervention in Kindergarten and First Grade

Working Paper
This version: February 2015

Authors and Affiliations:
Jade M. Jenkins[1] (presenting author)
Tyler W. Watts[1]
Katherine Magnuson[2]
Douglas Clements[3]
Julie Sarama[3]
Christopher B. Wolfe[4]
Mary Elaine Spitler[5]

[1] School of Education, University of California, Irvine
[2] School of Social Work, University of Wisconsin, Madison
[3] College of Education, University of Denver
[4] Indiana University-Kokomo
[5] University of Buffalo, State University of New York

Abstract

Little research has focused on why short-term gains from preschool may disappear and the conditions under which gains from preschool might be sustained into elementary school. We investigate whether two aspects of the elementary school environment may help to sustain the academic gains made during preschool using two random assignment preschool studies: 1) whether advanced and challenging instruction in kindergarten and first grade; 2) professional supports in which preschool teachers interact with their kindergarten and first grade counterparts to coordinate instruction and transition. We also assess whether the child's home learning environment moderates the persistence of preschool effects. We did not find any evidence to support the hypothesis that better instructional quality mitigates the fadeout of preschool treatment effects during elementary school. However, we did find some evidence that when the preschool intervention was coupled with teacher professional supports in kindergarten and first grade, this all but eliminated the fadeout of effects observed between kindergarten and first grade. We also did not find that factors in the home environment, parents education and home learning activities, help to sustain the gains made during preschool. Future research should investigate aligned preschool-elementary school curricular approaches to sustain the benefits of ECE programs for low-income children.

Introduction

A substantial literature documents the benefits of early childhood education and formal preschool experiences on children's school readiness, with low-income and otherwise disadvantaged children benefitting the most from these programs (Barnett, 2011; Camilli, Vargas, Ryan, & Barnett, 2010; Duncan & Magnuson, 2013; Reynolds, Temple, & Ou, 2010). However, these academic benefits often fade as children age, and most disappear by the end of kindergarten or first grade (Barnett, 1995; Currie, 2001; Puma, Bell, Cook, & Heid, 2010). A meta-analytic study estimated the magnitude of preschool intervention fadeout at .025 standard deviations per post-treatment year (Leak et al., 2013).

Little research has focused on why short-term gains from preschool may disappear and the conditions under which gains from preschool might be sustained into elementary school (Claessens, Engel, & Curran, 2013). One hypothesis of preschool fade out is that children's elementary school teachers continue to teach content that children already learned during preschool, thus curtailing academic growth. Indeed, recent work suggests that spending too much instructional time on content already mastered by students may temper achievement gains, whereas exposure to more advanced content in kindergarten could bolster new skill development (Engel, Claessens, & Finch, 2013; Magnuson, Ruhm, & Waldfogel, 2007). Another hypothesis is that the short-term gains from preschool are best maintained if the child's home environment continues to provide supportive interactions and promotes the academic skills children acquire during preschool.

In the current study, we investigate two salient approaches available to policymakers that may improve preschool participants' instructional experiences in elementary school. The first involves advanced and high-quality instruction in kindergarten and first grade, because children who attend preschool will hypothetically benefit more from more rigorous content. The other involves some type of professional support in which preschool teachers interact with their kindergarten and first grade counterparts to develop a seamless transition from one grade to the next. We use two experimental studies of preschool interventions and children's elementary school environments to examine whether the quality of instructional content or providing professional development supports to early grade teachers moderate the impacts of two well-known programs on children's cognitive skills: Head Start and Building Blocks. We also assess whether the child's home learning environment, as measured by learning activities in the home and parent's education, moderate the persistence of preschool program effects on children's intermediate-term cognitive outcomes.

Background

Research from neuroscience, education, psychology, and economics compose much of the foundational literature for the importance of early childhood education (ECE) programs and the role of early intervention in preventing long-term problems—particularly for children in poverty (Barnett, 2011; Bowman, Donovan, & Burns, 2000; Duncan, Ziol-Guest, & Kalil, 2010; Fox, Levitt, & Nelson Iii, 2010; Hackman & Farah, 2009; Knudsen, 2004; Magnuson & Waldfogel, 2005; McLoyd, 1998; Peisner-Feinberg et al., 2001; Shonkoff & Meisels, 2000; Trachtenberg & Stryker, 2001). In the short-term, high-quality early childhood education programs provided to children from at-risk groups improve children's cognitive and language development (Yoshikawa et al., 2013). However, evidence from a recent meta-analysis suggests that these effects almost always fade during the early elementary school years (Leak et al., 2013).

Nevertheless, researchers and practitioners continue to target early childhood as a particularly opportune time for investment, with much of this continued support for ECE programs influenced by the findings of two well-known interventions: Perry Preschool and Abecedarian. These evaluations showed that intensive (and expensive) early education programs for disadvantaged children can improve cognitive and language abilities by .75 to 1.5 standard deviations at the end of treatment. In the long-term, Abecedarian participants had lower levels of grade retention and high school dropout, reduced placement in special education, and higher rates of attending a 4-year college and having a full-time job as a result of treatment (Barnett & Masse, 2007; Campbell et al., 2012; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Campbell et al., 2008). For Perry Preschool participants, the program produced lasting effects through age forty on employment, earnings, and reduced the likelihood of adult poverty rates and arrest (Belfield, Nores, Barnett, & Schweinhart, 2006; Schweinhart, 2005).

Based on this literature, the U.S. has seen an unprecedented expansion of ECE programs over the past 40 years, primarily in the form of state pre-kindergarten programs (Jenkins, 2014). Yet, evidence of positive long-run impacts from newer programs remains nearly non-existent. When compared with Perry Preschool and Abecedarian, evaluations of these newer programs differ in a few key ways. First, modern ECE programs are typically much more modest in scope, as they usually involve one year of center-based care. Second, comparison groups in evaluation studies now have access to far more forms of alternative care, thus program participation in the treatment group is often compared with a control group that includes children participating in other, possibly high-quality, early childhood programs. As such, experimental evaluation of the one-year state pre-k program in Tennessee indicates that the short-term gains made during the preschool year disappear completely in elementary school (Lipsey, Hofer, Dong, Farran, & Bilbrey, 2013). A recent nonexperimental evaluation (using propensity scores) of Oklahoma's pre-k program indicate that the large end-of-treatment effects disappeared by third-grade, with effects remaining only for boys math skills (Hillm, Gormley, & Adelstein, 2015).

However, some studies do find evidence of persistent preschool effects. Exploiting variation in the timing of Texas pre-k implementation across districts, Andrews et al. show boosts to third grade reading and math scores and reduced special Education and retention rates (2012). Using similar variation in between-county program rollout and expenditures, Ladd and colleagues found that North Carolina's preschool investments improve aggregate educational outcomes at third grade (Ladd, Muschkin, & Dodge, 2014; Muschkin, Ladd, & Dodge, 2015).

The Head Start Impact Study experiment (HSIS) is perhaps infamous for showing impacts at the end of treatment and fadeout in elementary school. In 2002, two cohorts of children were randomly assigned to receive Head Start services at sites across the country. The end-of-program-year effect sizes average 0.2 SD for both the age-3 and age-4 cohorts on early language and literacy skills, and a .15 SD effect size on early math skills for age-3 cohort participants (Puma et al., 2010). Subsequent analyses show that these effects also vary by counterfactual conditions (i.e., different treatment effect estimates depending on whether control children attended quality preschool; Feller, Grindal, Miratrix, & Page, 2014; Walters, 2014). However, the short-term gains from HS disappear in the study's follow-up of in children's kindergarten, first, and third grade years (Puma et al., 2012).

On the other hand, analyses of Head Start participants in other strong quasi-experimental studies find substantial long-run effects, with positive impacts on academic and health outcomes of 0.2-0.3 standard deviations (Currie & Thomas, 1995; Deming, 2009; Garces, Thomas, & Currie, 2002; Ludwig & Phillips, 2008). Recent evidence suggests that these long-run effects are

maintained through children's improved social and emotional or personality skills (Chetty et al., 2011; Heckman, Pinto, & Savelyev, 2013; Love, Chazan-Cohen, Raikes, & Brooks-Gunn, 2013). As such, researchers in this area are trying to reconcile the disappearance of short-term academic effect with observed long-term benefits. Often discussed in the literature are children's subsequent developmental contexts, and their role in sustaining or erasing the impacts of preschool. For young children, these contexts are their elementary schools and their home and family environments.

School Environment

One common explanation for differences in the persistence of the impact of early education programs focuses on children's subsequent school experiences. Disadvantaged children and children who live in low-income neighborhoods—those targeted by public ECE interventions—are more likely to enter lower performing schools with lower quality instruction (M. A. Clements, Reynolds, & Hickey, 2004; Crosnoe & Cooper, 2010; McLoyd, 1998; Pianta, Belsky, Houts, & Morrison, 2007). When children leave quality preschools to attend low-quality schools, they have fewer opportunities for maintaining the benefits of preschool across the early grades and lose their early developmental gains (Currie & Thomas, 2000; V. E. Lee & Loeb, 1995; Reynolds, Ou, & Topitzes, 2004; Zhai, Raver, & Jones, 2012).

It seems intuitive that enriching instructional experiences in early elementary school are critical for the academic skills of preschool graduates; without them, preschool participants could lose their early advantage as their classmates slowly "catch up" to their skill level (Barnett, 2011; McKey, 1985; Zigler & Styfco, 2004). Yet, measuring and understanding what a high-quality elementary experience would be for preschool graduates is not obvious. Furthermore, children who do not attend preschool may benefit equally as much as preschool participants do from high-quality education. For example, evidence suggests that smaller class sizes improve children's long-run outcomes (Chetty et al., 2011; Nye, Hedges, & Konstantopoulos, 2000). However, Magnuson et al. found that that smaller class sizes in kindergarten had a greater impact on children who did *not* attend preschool. When children experienced smaller class sizes as well as more instruction time, their peers who did not attend preschool were able to catch up to their level of achievement more rapidly (Magnuson et al., 2007). Others have suggested that more instruction and experiencing a full-day rather than part-day kindergarten is associated with improved short-run academic outcomes (Cannon, Jacknowitz, & Painter, 2011; Gibbs, 2014; V. Lee, Burkam, Ready, Honigman, & Meisels, 2006), but other findings indicate that any benefits from a full-day may diminish over time (Votruba-Drzal, Li-Grining, & Maldonado-Carreño, 2008).

There is also evidence that children receive instruction that is misaligned with the instruction and skill they received during preschool. Recent work suggests that spending too much instructional time on content already mastered by preschool graduates may temper achievement gains, whereas exposure to more advanced content in kindergarten could bolster new skill development (Engel et al., 2013; Engel, Claessens, Watts, & Farkas, 2014; Magnuson et al., 2007). Furthermore, advanced content is beneficial for all kindergarten students, regardless of whether they attended preschool (Claessens et al., 2013).

A high quality instructional experience in elementary school may involve alignment between what was learned in pre-k and what is taught in kindergarten and first grade. Indeed, researchers advocate for an integrated "Preschool to grade three" or PK-3 approach to education (Bogard & Tananishi, 2005; Kagan & Kauerz, 2012). This involves the vertical alignment (i.e.,

across grades) of teacher quality, instructional tools and curriculum to create an overall set of educational experiences that build on one another. Recently, the Federal Race to the Top Early Learning Challenge grants identified "creating preschool through third grade approaches to sustain improved early learning outcomes through the elementary grades" as an essential component for the newest round of funding.

Still, there is limited empirical evidence on what this approach would need to look like to be effective. The Chicago Child Parent Center (CPC) program used a PK-3 approach when preschool participants moved into kindergarten that involved curricular supports, parent involvement, and school-wide services like health and links to community resources. Using nonexperimental analyses, they found that children in the CPC group were more likely to finish high school, completed more years of education, and had lower rates of juvenile arrest and school dropout relative to comparison children (Reynolds & Temple, 1998; Reynolds et al., 2010; Reynolds, Temple, Robertson, & Mann, 2001, 2002).

Another approach could involve helping teachers design curriculum that does not repeat information that students already learned during their ECE program. Clements and colleges (2013) evaluated the scale-up of a highly successful preschool mathematics curriculum in state preschool programs serving low-income communities in New York and Massachusetts. They found that random assignment to the preschool mathematics intervention had a large impact on end-of-preschool mathematics scores, but this effect faded substantially by first grade (D. Clements, Sarama, Wolfe, & Spitler, 2013). However, some students were assigned to another condition that included the preschool mathematics intervention, but also featured additional pedagogical development sessions for their kindergarten and first grade teachers. In these sessions, teachers were informed of the mathematics content the students had learned in preschool, with the hope that teachers would not spend too much time repeating this information in kindergarten and first grade. Results showed that when compared with children who only received the preschool intervention, students assigned to the follow-through condition had substantially less-effect fade out at the end of first grade.

In sum, these findings highlight that the target population of many ECE programs may spend their first year in public school in classrooms that may not capitalize on the gains made during preschool. For ECE programs and policies to make a substantial impact on the well-being of young, disadvantaged children, it is critical to understand how to successfully build on investments in early learning when children enter elementary school. We do this in our study.

Home environment

Though fadeout research has focused primarily on the instructional environment, a child's home environment can help to maintain school readiness and growth in academic skills. In their foundational work, Coleman et al. stress that the family environment is central in children's academic success (1966). Subsequent work confirms the hypothesis that the early childhood home environment accounts for much of the early achievement gaps which lead to the wider gaps present in subsequent years (Fryer & Levitt, 2006; Yeung & Pfeiffer, 2009). Exposure to stimulating activities in the home such as reading, counting, and new vocabulary growth are all associated with academic success (Bradley, McKelvey, & Whiteside-Mansell, 2011; Brooks-Gunn & Markham, 2005; Crosnoe & Cooper, 2010; Melhuish et al., 2008). Therefore, children who have more home learning opportunities may be better able to maintain the benefits of preschool through their elementary school years.

Parents' education is particularly important in this regard. Because parental education is

associated with complex language stimulation, children in homes with educated parents have had 50 percent more exposure to language as children of less-educated parents by the age of three (Hart & Risley, 1995). Parents' education also has the greatest impact on the amount of time and quality of time spent with children, as well as on other investments that parents make in children's development (Guryan, Hurst, & Kearney, 2008; Hoff, 2013; Kalil, 2015; Kalil, Ryan, & Corey, 2012; Leibowitz, 1977; Ramey & Ramey, 2009). Educated mothers also benefit their child's development through cultural knowledge and social connections (Harding, Morris, & Hughes, 2015).

A rather obvious point also tempers these findings; even if preschool generates meaningful benefits, children are still embedded in the context of poverty. These continued stressful or negative experiences due to low income and associated risk factors would work against the early learning gains from ECE. For this reason, it may be that time-limited ECE programs on their own cannot eliminate the detrimental effects of poverty on children's cognitive or social-emotional development across the lifespan (Brooks-Gunn, 2003). Still, investigating whether and how these factors influence the persistence of preschool effects is essential for policy.

Present Study

Although it has been suggested that the persistence of short-term preschool impacts on cognitive skills are shaped by students' subsequent classroom and home experiences, a limited number of studies have tested these hypotheses directly using random assignment to preschool treatment. Our analyses assess the extent to which the persistence of preschool program effects on children's cognitive skills depends upon the features of the kindergarten and first grade classroom they attend in two preschool studies, the Head Start Impact Study and the Building Blocks Preschool Mathematics intervention study. We identified two key instructional characteristics in both studies to operationalize the quality of elementary school exposure to literacy content (HSIS) and math content (BB). We also assess whether the extent to which the child's family promotes learning in the home moderates the persistence of preschool effects.

The primary research questions for this study are:

1. Does the quality of academic instruction in kindergarten and first grade moderate the magnitude of preschool intervention effects on children's academic skills in kindergarten and first grade?
    a. Does a professional development intervention for kindergarten and first grade teachers that provided techniques designed to build upon the preschool program moderate preschool intervention effects on children's academic skills in kindergarten and first grade?
2. Does parental education or the amount of learning activities in the home moderate the magnitude of preschool intervention effects on children's academic skills in kindergarten and first grade?

Note that treatment was randomly assigned in the interventions we study, but the hypothesized moderators were not. The majority of our results are associational and make the (strong) assumption that classroom instructional quality and home learning activities are distributed randomly across treatment and control groups and across children's potential outcomes.

Preschool Interventions. Our study uses data from two preschool interventions—Head Start and the Building Blocks preschool mathematics intervention. *Head Start* is a comprehensive child development program that provides children with preschool education, health screenings and examinations, and nutritious meals, in a full-day, center-based setting. The Head Start children in our sample participated in the program during their pre-kindergarten year at age 4 at different research sites across the country. The Head Start Impact Study (HSIS) evaluation began in 2002 (described below).

Building Blocks (BB) is a preschool mathematics curriculum that encourages the acquisition of conceptual and procedural knowledge in both numeracy and geometric/spatial reasoning through the emphasis of empirically-supported learning trajectories (see D. Clements & Sarama, 2008). The TRIAD (Technology-enhanced, Research-based, Assessment, and professional Development) evaluation study was designed to assess the long-run impacts of BB in 42 public elementary schools operating state preschool programs serving low-income communities in Boston, Massachusetts and Buffalo, New York. Study schools were assigned to one of three conditions: 1) BB preschool curriculum; 2) BB preschool curriculum with follow-through; 3) control (business as usual). Children in schools assigned to the two BB groups received the BB curriculum during preschool (age 4), and preschool teachers attended 13 study-administered pedagogical development (PD) sessions throughout the preschool year. Teachers in schools assigned to the "BB with follow-through" group received additional PD designed to help bridge the gaps between preschool, kindergarten, and first grade. These additional PD sessions brought teachers from all three grades together to discuss what students learn in each grade, and minimize the amount of repeated content.

Data

Our use of two different random assignment studies allows us to test the robustness of our findings. However, the family, child, and classroom characteristics differ slightly between each dataset. We have aligned our variables so that they are as similar as possible across datasets, but we present our data, analyses and results separately by study for ease of exposition.

Preschool Intervention I: Head Start

The Head Start sample comes from the HSIS experiment dataset, which is a nationally representative sample of Head Start participants and a group of comparable non-participants. The full sample includes newly entering 3-and 4-year old Head Start applicants who were randomly assigned to receive the Head Start program or a control group that did not enroll in Head Start where parents either found other available services for their child or the child was cared for at the home. Baseline survey and child assessment data were collected by study investigators (Westat) in the Fall of 2002, at post-treatment child assessments were collected at the end of Head Start in Spring 2003, and during kindergarten and first grade in Spring 2004 and 2005. *Our analyses use the 4-year-old cohort* only so that the children in both of HSIS and BB analyses received the preschool intervention during the same developmental period. This sample was further limited to children whose teachers responded to the survey (n=1080).

Table 1 presents descriptive statistics for the current analytic sample. The children and families in the sample are all very low income and have the following characteristics: 45% Hispanic, 39% white and 15% Black, 42% of parents have less than a high school degree, 23% are recent immigrants, 16% are teenage mothers and a majority (84%) live in an urban area. Information on children's elementary school experiences were collected from kindergarten and first grade teachers through a teacher survey in the spring of 2004 and 2005. Literacy skills were

measured with the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997) and Letter Word and Spelling standard scores from the Woodcock-Johnson Psycho-Educational Battery-Revised III (Woodcock, McGrew, & Mather, 2001). We created a literacy assessment composite measure to use as the dependent variable by standardizing all three measures to mean 0 and standard deviation of 1, averaging across the three, and then restandardizing the measure.

Variables of interest.

Classroom environment. In the HSIS, teachers were asked how many times in the past week their class did a given literacy activity. We coded each activity into basic or advanced based on grade level (available in Appendix A). We converted each basic and advanced activity into times per month by taking the mean value of the answer category (e.g., Never=0, 1-2 times per week=1.5), multiplied by 4, and then standardized this measure to have a mean of 0 and standard deviation of 1. Instructional quality during the first grade year is a cumulative measure of quality from both Kindergarten and first grade, averaged across the two years.

Home environment. We identified two sets of home learning activities based on items asked in the parent survey at the end of the HSIS treatment year; home literacy activities and other general home learning activities (items used in each index available in Appendix B). The home literacy items were answered on a 1-4 Likert scale. We averaged this number across all items the and then standardized the measure to have a mean of 0 and standard deviation of 1. The general home learning activities were answered as 1= true or 0=not true. We added these items together and standardized the measure to mean 0, standard deviation of 1. We also include an indicator of parent's education, having greater than a high school degree, to test for moderation.

Preschool Intervention II: Building Blocks

TRIAD study participants (n=1375) were randomly selected from study schools at the beginning of the preschool year (2006-2007). The current study includes two different analytic samples. The first sample consists of students who had at least one non-missing classroom observation measure, and had complete data on achievement tests in preschool, kindergarten and first grade (n= 865). The second analytic sample includes only students who had valid test score data in preschool and kindergarten, and non-missing parent survey data (n=887).

Table 1 presents descriptive statistics for the full sample, as well as p-values indicating whether assignment to any of the study conditions was correlated with each of the respective baseline characteristics. In the full sample, 35% of students were assigned to the BB group and 36% were assigned to the BB with follow-through group. The majority of students qualified for free or reduced price lunch (84%); 53% identified as African American and 22% as Hispanic. Math achievement was assessed at preschool entry, and at the end of the preschool, kindergarten and first grade year via the Research Based Early Mathematics Assessment (REMA; D. Clements, Sarama, & Liu, 2008). Further, we did not find any baseline characteristic to be unbalanced between the three study conditions.

Students included in the first analytic sample were less likely to be male (6%, $p < .05$), and more likely to be free or reduced price lunch (30%, $p < .001$). Students included in the second analytic sample were less likely to be Hispanic (6%, $p < .05$), designated as limited English proficient (7%, $p < .05$), and were slightly younger at preschool entry (.06 years, $p < .05$).

Variables of interest.

Classroom environment. In the TRIAD evaluation of BB, teachers' instructional practices were evaluated via the Classroom Observation of Early Mathematics Environment and Teaching (COEMET; see D. Clements, Sarama, Spitler, Lange, & Wolfe, 2011). The COEMET is composed of 28 Likert-scaled items. Assessors, who were blind to treatment group, rated classrooms for teaching practices known to support early math development, such as the use of engaging small group activities and emphasizing cognitively demanding concepts and strategies. For the kindergarten year, we took the average of these 28 items and then standardized scores. As with the HSIS, our measure of first grade instructional quality is the standardized average of a child's kindergarten and first grade COEMET scores. We also included the number of mathematical activities observed during each COEMET period in our analysis as an indicator of the amount of time spent on mathematics in the class.

Home environment. Parents of children participating in the evaluation were given a survey in which they were asked questions regarding time spent on home instructional activities. Items included questions regarding time spent practicing letters, words and numbers, among other topics. These 11 questions were scaled 0-3, with scores of 3 indicating that the parent and child spent time on the activity every day of the week, and scores of 0 indicating that they never participated in the activity. Parents were also asked about the number of children's books in the home, and this question was also coded 0-3, with a score of 3 representing over 100 books in home. We then averaged the 11 activity items together with the number of books item to create an index of home academic environment. This index was then standardized across treatment groups to a mean of 0 and standard deviation of 1. As with the HSIS, we also included a bivariate measure of parent education, with scores of 1 indicating at least come college education.

Methods

We use multivariate regression to estimate the effect of instructional experiences on the magnitude of preschool treatment effects in children's kindergarten and first grade year. Both preschool interventions were randomly assigned, so treatment effects estimated during the pre-k year and after are unbiased. In the HSIS models, we focus on literacy outcomes as the dependent variable, and in BB models, we focus on mathematics. In all models, we regressed achievement measures (taken at either end of preschool, kindergarten, or first grade) on treatment status, fixed effects for unit of random assignment, baseline assessment scores, and a set of control variables, varying slightly between BB and HSIS. We then add measures of classroom instruction as covariates to see how much of the treatment effect is explained by high-quality instructional practices. Finally, we add models in which treatment is interacted with classroom instruction. If high-quality instruction in kindergarten and first grade helps reduce fade-out, then these interactions should be positive and significant.

The HSIS was a very comprehensive study, and the dataset includes other characteristics about the kindergarten classroom environment, such as class size and proportion of children in poverty. Therefore we also include additional tests for moderation by classroom environment for the HSIS only, presented in the appendix. Notably, we were able to conduct a kindergarten classroom fixed-effect model, comparing outcomes of treatment and control children experiencing the same instructional environment. We include these models to test for additional mechanisms hypothesized in prior research about the suppression or maintenance of treatment effects in elementary school.

To test for moderation by factors in the home environment, we followed the same

analytic strategy outlined above but instead add our indices of home literacy (HSIS) and home academic (BB) activities, and maternal education level, and then interact these variables with treatment. Because the BB study only included measures of home activities during the preschool year, we only conduct analyses for moderation by the home environment for outcomes at the end of the kindergarten year.

To address differential attrition and imbalances between treatment and control children at baseline in the HSIS, we follow Bitler et al. (2014) use inverse probability of treatment weights for all HSIS models. These weights effectively control for baseline characteristics shown in Table 1. Note that these weights do not adjust for Kindergarten and first grade *teacher* nonresponse. Analyses of outcomes based on teacher response status available in Appendix C.

Results

Table 1 presents the descriptive statistics for both the HSIS and BB samples. All HS results are presented in Table 2, and all BB results are presented in Table 3. All variables except the treatment indicator and parents' education are scaled in standard deviation units to facilitate their interpretation as effect sizes. Note that models testing for moderation from the home environment were conducted for the Kindergarten year only, whereas moderation from the classroom environments were conducted for both Kindergarten and first grade.

Preschool Intervention I: Head Start

Kindergarten.

Classroom environment. Table 2a presents the results for moderation of impacts on literacy and language skills for the HSIS sample, where the dependent variable is a composite of three literacy and language assessments. Model 1 shows the end of Head Start year treatment effect, which is significant with an effect size of .16. However, this effect becomes negative and insignificant by the end of Kindergarten (Model 2). When we add the instructional quality variables (basic and advanced literacy instruction) in model 3, the treatment effect remains unchanged though the coefficient on advanced literacy activities is .12 and significant, and the coefficient on basic literacy activities is -.12 and significant. Model 4 adds the interactions between instructional quality and treatment. Neither term was statistically significant, suggesting that advanced literacy instruction is unable to sustain the gains of the Head Start treatment group children through the kindergarten year.

Appendix D presents the results for similar models that test for moderating effects of attending full-day kindergarten, kindergarten class size, and the classroom-level proportion of children in poverty (as measured by Free and Reduced-price lunch). None of these variables, when interacted with treatment, were statistically significant. We were also able to estimate a kindergarten classroom fixed effect model (Model 1), which tests whether Head Start participants have stronger literacy skills in kindergarten relative to a control child in the same kindergarten classroom (~250 children shared a classroom with a control child). The treatment effect in the kindergarten fixed effect model was not significant, was small in magnitude (.02).

Home environment. Models 1-5 in Table 2b show the effect of Head Start on our literacy composite at the end of Kindergarten. Model 1 shows the end of Head Start year treatment effect, and Model 2 shows the treatment effect at the end of kindergarten for children whose parents responded to the end of Kindergarten survey, which is negative and insignificant. Model 3 adds parents' education (1=> than High School degree) and the home literacy activities index score, both of which are positive and significant (.43 and .15, respectively). Model 4 adds the

interactions between the home literacy activities and treatment, and parents' education and treatment. Neither term was statistically significant, suggesting that home learning experiences and parent's education are not able to sustain the gains of Head Start treatment group children through the kindergarten year. In Model 5 we replace the home literacy activities index with the general home learning activities index, and see the same pattern of results.

First Grade.

Classroom environment. Models 1-4 in Table 2c show the effect of Head Start on our literacy composite at the end of first grade. In each model, the treatment effect is not significant. Including the instructional quality variables in Model 3 does not change the significance or magnitude of the coefficients. Neither the instructional variables nor the interactions with treatment were significant in Model 4 as well. Again, there is no evidence that high-quality instruction can sustain Head Start gains through the first-grade year.

Preschool Intervention II: Building Blocks

Kindergarten.

Classroom environment. In Table 3a, Models 1-6 display the impacts of BB on kindergarten mathematics achievement. Model 1 shows the BB treatment effect at the end of the preschool year with an effect size of .66. At Kindergarten (model 2) the effect drops to .34 and remains significant. When we add the instructional quality variables (COEMET and number of math activities) in Model 3, the treatment effect remains unchanged, and the coefficient on the number of math activities is .13 and significant ($p < .04$). Model 4 adds the interactions between instructional quality and treatment, but neither term was significant.

Models 5 and 6 take a different look at sustaining pre-K gains. Rather than classroom quality, they examine a kindergarten teacher PD focused directly on sustaining the BB preschool gains. Unlike our measures of teacher-driven classroom instructional quality, teachers were randomly assigned to engage in additional PD. The treatment effect for students in the "treatment with follow-through" was .38 and significant, but it was not significantly different from the end of kindergarten impact for students who received BB without follow-through PD (.32). All told, there is no evidence that either instructional quality or focused PD is able to sustain pre-K gains through the end of kindergarten.

Home environment. Models 1-4 of Table 3b show the effect of high academic stimulation in the home environment on curbing fade out effects through kindergarten. As with the classroom environment models, the treatment effect was reduced by nearly half between the end of preschool and spring of kindergarten. In Model 3, we add our measure of home academic activities and mother's education to the model, and find not mediation of the treatment effect at the end of kindergarten, though the main effect of the mother having at least some college education was marginally significant (.12, $p < .10$). Model 4 presents the results of a model in which treatment status was interacted with our measures of home environment, and neither interaction was statistically significant. We also tested interactions between these home measures and assignment to the follow-through condition (not shown), and these interactions also failed to reach statistical significance.

First Grade. In Table 3c, we present results from models examining the relation between high-quality mathematics instruction in kindergarten and first grade and treatment effects at the end of first grade. Model 1 again presents the large treatment effect at the end of preschool, and Model 2 displays that this effect dropped to marginally significant .15 ($p < .10$). When we add the instructional quality variables in Model 3, the BB only treatment effect falls

further to .11, and found a significant positive effect for the number of math activities. Adding the interaction with treatment in Model 4 produces a similar pattern, where the BB only treatment effect is not significant at the .05 level, but there is a .21 significant ($p < .05$) effect size for the number of math activities. The interaction between instructional quality and treatment was not significant. Thus, yet another test shows no evidence that high-quality instruction sustains pre-K gains.

As with the final models in Table 3b, Model 5 in Table 2 focuses on a PD approach to sustaining gains. A comparison of treatment effects for both the BB only and the BB plus follow-up group shows that the effect size for the follow-up group is .32 and significant ($p < .001$), compared with a .18 effect size ($p < .05$) for the BB only group. The .32 coefficient is only slightly smaller than the .37 coefficient found on the PD group at the end of kindergarten, suggesting very little fadeout during first grade. In other words, two years of PD may be able to help sustain gains, although a comparison of the two effect sizes revealed that the follow-through group effect is not quite statistically significantly larger than the preschool only effect ($F= 3.14$; $p = .08$).

Selection into classroom environments

A primary concern with the moderation terms we estimate for both BB and HSIS is that children and families may select into different types of classroom environments post-treatment. Assuming that children with better potential outcomes would me more likely to experience sustaining environments in Kindergarten and first grade, the estimates presented above are upwardly biased and represent an upper bound of the true effect of the environmental condition. We test for selection into classroom environments explicitly using the HSIS, which included more extensive classroom and school characteristics than the BB study, shown in Appendix E. Here we regress a number of Kindergarten classroom and school characteristics (e.g., class size, teacher's education, school reading proficiency level) on children's treatment status, including fixed effects for center of random assignment and the IPT weights. These tests reveal that at least correlationally, there is no evidence of differential selection into sustaining environments by preschool treatment status.

Discussion

[TBD]

Limitations

While we use two preschool interventions where treatment was randomly assigned in the, but the classroom and home moderators were not. Children select into different types of elementary school environments, and these selection processes are not likely to be independent of academic outcomes at the end of Kindergarten and first grade. Therefore, perhaps with the exception of the HSIS kindergarten classroom fixed effect, our results are associational and make the strong assumption that classroom instructional quality and home learning activities are distributed randomly across treatment and control groups and across children's potential outcomes.

Conclusion

We did not find any evidence to support the hypothesis that better instructional quality mitigates the fadeout of preschool treatment effects during elementary school. However, we did find some evidence that when the BB intervention was coupled with teacher professional

supports in kindergarten and first grade, this all but eliminated the fadeout of effects observed between kindergarten and first grade. Still, with both the focused PD and high instructional quality, BB could not reduce fadeout effects between preschool and kindergarten. We also did not find that factors in the home environment, parents education and home learning activities, help to sustain the gains made during preschool. Future research should investigate aligned preschool-elementary school curricular approaches to sustain the benefits of ECE programs for low-income children.

References

Andrews, R. J., Jargowsky, P., & Kuhne, K. (2012). The Effects of Texas's Targeted Pre-Kindergarten Program on Academic Performance. National Bureau of Economic Research Working Paper Series, No. 18598.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. Future of Children, 5, 25-50.

Barnett, W. S. (2011). Effectiveness of Early Educational Intervention. Science, 333(6045), 975-978.

Barnett, W. S., & Masse, L. N. (2007). Comparative benefit-cost analysis of the Abecedarian program and its policy implications. The Economics of Early Childhood Education, 26(1), 113-125.

Belfield, C. R., Nores, M., Barnett, W. S., & Schweinhart, L. (2006). The High/Scope Perry Preschool Program. Journal of Human Resources, XLI(1), 162-190.

Bitler, M. P., Hoynes, H. W., & Domina, T. (2014). Experimental Evidence on Distributional Effects of Head Start. National Bureau of Economic Research Working Paper Series, No. 20434.

Bogard, K., & Tananishi, R. (2005). Pk-3: An aligned and coordinated approach to education for children 3 to 8 years old. Social Policy Report, XIX(III), 1-24.

Bowman, B. T., Donovan, S., & Burns, M. S. (2000). Eager to learn: educating our preschoolers By N: National Research Council (U.S.). Committee on Early Childhood Pedagogy.

Bradley, R. H., McKelvey, L. M., & Whiteside-Mansell, L. (2011). Does the Quality of Stimulation and Support in the Home Environment Moderate the Effect of Early Education Programs? Child Development, 82(6), 2110-2122.

Brooks-Gunn, J. (2003). Do you believe in magic? What we can expect from early childhood intervention programs. Child Development Social Policy Report, XVII(1).

Brooks-Gunn, J., & Markham, L. B. (2005). The Contribution of Parenting to Racial and Ethnic Gaps in School Readiness. The Future of Children, 15(1).

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. Teachers College Record, 112(3), 579-620.

Campbell, F. A., Pungello, E., Burchinal, M. R., Kainz, K., Pan, Y., Wasik, B. H., et al. (2012). Adult outcomes as a function of an early childhood educational program: an Abecedarian Project follow-up. Developmental Psychology, 48(4), 1033.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. Applied Developmental Science, 6(Journal Article), 42-57.

Campbell, F. A., Wasik, B. H., Pungello, E., Burchinal, M. R., Barbarin, O. A., Kainz, K., et al. (2008). Young adult outcomes of the Abecedarian and CARE early childhood educational interventions. Early Childhood Research Quarterly, 23(4), 452-466.

Cannon, J. S., Jacknowitz, A., & Painter, G. (2011). The effect of attending full-day kindergarten on English learner students. Journal of Policy Analysis and Management, 30(2), 287-309.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. The Quarterly Journal of Economics, 126(4), 1593-1660.

Claessens, A., Engel, M., & Curran, F. C. (2013). Academic Content, Student Learning, and the

Persistence of Preschool Effects. American Educational Research Journal.
Clements, D., & Sarama, J. (2008). Experimental Evaluation of the Effects of a Research-Based Preschool Mathematics Curriculum. American Educational Research Journal, 45(2), 443-494.
Clements, D., Sarama, J., & Liu, X. H. (2008). Development of a measure of early mathematics achievement using the Rasch model: the Research-Based Early Maths Assessment. Educational Psychology, 28(4), 457-482.
Clements, D., Sarama, J., Spitler, M. E., Lange, A. A., & Wolfe, C. B. (2011). Mathematics learned by young children in an intervention based on learning trajectories: A large-scale cluster randomized trial. Journal for Research in Mathematics Education, 42(2), 127-166.
Clements, D., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2013). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies Persistence of effects in the third year. American Educational Research Journal, 50(4), 812-850.
Clements, M. A., Reynolds, A. J., & Hickey, E. (2004). Site-level predictors of children's school and social competence in the Chicago Child–Parent Centers. Early Childhood Research Quarterly, 19(2), 273-296.
Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). Equality of educational opportunity. Washington, DC.
Crosnoe, R., & Cooper, C. E. (2010). Economically Disadvantaged Children's Transitions Into Elementary School: Linking Family Processes, School Contexts, and Educational Policy. American Educational Research Journal, 47(2), 258-291.
Currie, J. (2001). Early Childhood Education Programs. Journal of Economic Perspectives, 15(2), 213-238.
Currie, J., & Thomas, D. (1995). Does Head Start Make a Difference? The American Economic Review, 85(3), 341-364.
Currie, J., & Thomas, D. (2000). School Quality and the Longer-term Effects of Head Start. Journal of Human Resources, 35(4), 755-774.
Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. American Economic Journal: Applied Economics, 1(3), 111-134.
Duncan, G. J., & Magnuson, K. (2013). Investing in Preschool Programs. The Journal of Economic Perspectives, 27(2), 109-132.
Duncan, G. J., Ziol-Guest, K. M., & Kalil, A. (2010). Early-Childhood Poverty and Adult Attainment, Behavior, and Health. Child Development, 81(1), 306-325.
Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test--Third Edition (PPVT-III). Upper Saddle River, NJ: Pearson Publishing.
Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching Students What They Already Know? The (Mis)Alignment Between Mathematics Instructional Content and Student Knowledge in Kindergarten. Educational Evaluation and Policy Analysis, 35(2), 157-178.
Engel, M., Claessens, A., Watts, T. W., & Farkas, G. (2014). The misalignment of kindergarten mathematics content.Unpublished manuscript, Under review.
Feller, A., Grindal, T., Miratrix, L., & Page, L. (2014). Compared to what? Variation in the Impacts of Early Childhood Education by Alternative Care-Type Settings, Working Paper.
Fox, S. E., Levitt, P., & Nelson Iii, C. A. (2010). How the Timing and Quality of Early Experiences Influence the Development of Brain Architecture. Child Development,

81(1), 28-40.
Fryer, R. G., & Levitt, S. D. (2006). The Black-White Test Score Gap Through Third Grade. American law and economics review, 8(2), 249-281.
Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. The American Economic Review, 92, 999-1012.
Gibbs, C. R. (2014). Experimental evidence on early intervention: The impact of fullday kindergarten. Frank Batten School of Leadership and Public Policy Working Paper(2014-04).
Guryan, J., Hurst, E., & Kearney, M. (2008). Parental Education and Parental Time with children. Journal of Economic Perspectives, 22(3), 23-46.
Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. Trends in cognitive sciences, 13(2), 65-73.
Harding, J. F., Morris, P. A., & Hughes, D. (2015). The Relationship Between Maternal Education and Children's Academic Outcomes: A Theoretical Framework. Journal of Marriage and Family, 77(1), 60-76.
Hart, B., & Risley, T. (1995). Meaningful Differences in the Everyday Experiences of Young American Children. Baltimore, M.D.: Brookes.
Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. The American Economic Review, 103(6), 1-35.
Hillm, C. J., Gormley, W. T., & Adelstein, S. (2015). Do the short-term effects of a high-quality preschool program persist? Early Childhood Research Quarterly.
Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. Developmental Psychology, 49(1), 4.
Jenkins, J. M. (2014). Early childhood development as economic development: Considerations for state-level policy innovation and experimentation. Economic Development Quarterly, 28(1).
Kagan, S. L., & Kauerz, K. (Eds.). (2012). Early childhood systems: Transforming early learning. New York, NY: Teachers College Press.
Kalil, A. (2015). Inequality Begins at Home: The Role of Parenting in the Diverging Destinies of Rich and Poor Children. In P. R. Amato, A. Booth, S. M. McHale, & J. Van Hook (Eds.), Families in an Era of Increasing Inequality (Vol. 5, pp. 63-82): Springer International Publishing.
Kalil, A., Ryan, R., & Corey, M. (2012). Diverging Destinies: Maternal Education and the Developmental Gradient in Time With Children. Demography, 49(4), 1361-1383.
Knudsen, E. I. (2004). Sensitive Periods in the Development of the Brain and Behavior. Journal of cognitive neuroscience, 16(8), 1412-1425.
Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2014). From Birth to School: Early Childhood Initiatives and Third‐Grade Outcomes in North Carolina. Journal of Policy Analysis and Management, 33(1), 162-187.
Leak, J., Duncan, G. J., Li, W., Magnuson, K., Schindler, H., & Yoshikawa, H. (2013). Is Timing Everything? How Early Childhood Education Program Impacts Vary by Starting Age, Program Duration and Time Since the End of the Program, Working Paper: Center on the Developing Child, Harvard University.
Lee, V., Burkam, D. T., Ready, D. D., Honigman, J., & Meisels, S. J. (2006). Full‐Day versus

Half‐Day Kindergarten: In Which Program Do Children Learn More? American Journal of Education, 112(2), 163-208.
Lee, V. E., & Loeb, S. (1995). Where Do Head Start Attendees End Up? One Reason Why Preschool Effects Fade Out. Educational Evaluation and Policy Analysis, 17(1), 62-82.
Leibowitz, A. (1977). Parental Inputs and Children's Achievement. The Journal of human resources, 12(2), 242-251.
Lipsey, M. W., Hofer, K. G., Dong, N., Farran, D. C., & Bilbrey, C. (2013). Evaluation of the Tennessee Voluntary Prekindergarten program: Kindergarten and first grade follow-up results from the randomized control design. Nashville, TN: Vanderbilt University, Peabody Research Institute.
Love, J. M., Chazan-Cohen, R., Raikes, H., & Brooks-Gunn, J. (2013). What makes a difference: Early Head Start evaluation findings in a developmental context. Monographs of the Society for Research in Child Development, 78(1), vii-viii, 1-173.
Ludwig, J., & Phillips, D. A. (2008). Long-Term Effects of Head Start on Low-Income Children. Annals of the New York Academy of Sciences, 1136(Reducing the Impact of Poverty on Health and Human Development: Scientific Approaches), 257-268.
Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? Early Childhood Research Quarterly, 22(1), 18-38.
Magnuson, K. A., & Waldfogel, J. (2005). Early childhood care and education: Effects on ethnic and racial gaps in school readiness. The Future of Children, 12(1), 169-195.
McKey, R. H. (1985). The Impact of Head Start on Children, Families and Communities. Final Report of the Head Start Evaluation, Synthesis and Utilization Project.
McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. American Psychologist, 53(2), 185-204.
Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the Home Learning Environment and Preschool Center Experience upon Literacy and Numeracy Development in Early Primary School. Journal of Social Issues, 64(1), 95-114.
Muschkin, C. G., Ladd, H. F., & Dodge, K. A. (2015). Impact of North Carolina's Early Childhood Initiatives on Special Education Placements in Third Grade. Educational Evaluation and Policy Analysis, 0162373714559096.
Nye, B., Hedges, L. V., & Konstantopoulos, S. (2000). The Effects of Small Classes on Academic Achievement: The Results of the Tennessee Class Size Experiment. American Educational Research Journal, 37(1), 123-151.
Peisner-Feinberg, E., Burchinal, M. R., Clifford, R. M., Culkin, M., Howes, C., & Kagan, S. L. (2001). The relation of preschool child-care quality to children's cognitive and social development trajectories through second grade. Child Development, 72(Journal Article), 1534-1553.
Pianta, R. C., Belsky, J., Houts, R., & Morrison, F. (2007). Opportunities to Learn in America's Elementary Classrooms. Science (New York, N.Y.), 315(5820), 1795-1796.
Puma, M., Bell, S., Cook, R., & Heid, C. (2010). Head Start Impact Study. Final Report. Washington, DC.: U.S. Department of Health and Human Services, Administration for Children and Families.
Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., et al. (2012). Third grade follow-up to the Head Start Impact Study: Final Report. Washington, DC.

Ramey, G., & Ramey, V. A. (2009). The Rug Rat Race. National Bureau of Economic Research Working Paper Series, No. 15284.
Reynolds, A. J., Ou, S., & Topitzes, D. (2004). Path of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago Child-Parent Centers. Child Development, 75, 1299-1328.
Reynolds, A. J., & Temple, A. J. (1998). Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago Longitudinal Study. Child Development, 69, 231-246.
Reynolds, A. J., Temple, J. A., & Ou, S. (2010). Preschool education, educational attainment, and crime prevention: Contributions of cognitive and non-cognitive skills. Children and Youth Services Review, 32(8), 1054-1063.
Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2001). Long-term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest: A 15-Year Follow-up of Low-Income Children in Public Schools. JAMA : the journal of the American Medical Association, 285(18), 2339-2346.
Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2002). Age 21 Cost-Benefit Analysis of the Title I Chicago Child-Parent Centers. Educational Evaluation and Policy Analysis, 24(4), 267-303.
Schweinhart, L. J. (2005). Lifetime Effects: The High/Scope Perry Preschool Study through Age 40. Ypsilanti, M.I.: High/Scope Educational Research Foundation.
Shonkoff, J. P., & Meisels, S. J. (Eds.). (2000). The Handbook of Early Intervention (Second ed.). New York: Cambridge University Press.
Trachtenberg, J. T., & Stryker, M. P. (2001). Rapid Anatomical Plasticity of Horizontal Connections in the Developing Visual Cortex. J. Neurosci., 21(10), 3476-3482.
Votruba-Drzal, E., Li-Grining, C. P., & Maldonado-Carreño, C. (2008). A Developmental Perspective on Full- Versus Part-Day Kindergarten and Children's Academic Trajectories Through Fifth Grade. Child Development, 79(4), 957-978.
Walters, C. (2014). Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start. National Bureau of Economic Research Working Paper Series, No. 20639.
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson-III Tests of Achievement. Itasca, IL.
Yeung, W.-J. J., & Pfeiffer, K. M. (2009). The black-white test score gap and early home environment. Social science research, 38(2), 412-437.
Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W., et al. (2013). Investing in our future: The evidence base on preschool education. New York, NY: Foundation for Child Development, Society for Research in Child Development.
Zhai, F., Raver, C. C., & Jones, S. M. (2012). Academic performance of subsequent schools and impacts of early interventions: Evidence from a randomized controlled trial in Head Start settings. Children and Youth Services Review, 34(5), 946-954.
Zigler, E. F., & Styfco, S. J. (Eds.). (2004). The Head Start Debates. Maryland: Paul H. Brookes Publishing.

Table 1. Descriptive statistics for the Head Start Impact Study and Building Blocks samples

	mean	sd	N	p-level
Head Start Impact Study: Age 4 cohort				
Treatment	0.62		1530	
Control	0.38		1530	
Preschool Entry Literacy skills composite score (std.)	-0.01	1.04	1460	
Male	0.51		1530	
Black	0.22		1530	
Hispanic	0.43		1530	
Mothers Education				
Below High School	0.42		1530	
High School degree	0.31		1530	
> High School deg.	0.26		1530	
Limited English Prof.	0.32		1530	
Child has special needs	0.14		1530	
Building Blocks Study				
PreK Only Treatment Group	0.35		1375	
Follow-Through Treatment Group	0.36		1375	
PreK Entry Math Composite Score	-3.22	0.82	1305	0.72
Site 1	0.72		1375	0.95
Male	0.49		1375	0.66
Afrian American	0.53		1375	0.77
Hispanic	0.22		1375	0.76
Ethnicity- Other	0.06		1375	0.31
Age at PreK Entry	4.34	0.35	1305	0.38
Mother's Education				
Below High School	0.15	1049	1049	0.93
High School degree	0.32	1049	1049	0.90
> High School	0.53	1049	1049	0.98
Free/Reduced Lunch	0.84	1077	1077	0.37
Limited Eng Prof.	0.16	1375	1375	0.51
Special Education	0.17	1305	1305	0.87

Note. The p-level column presents the results from an F-test that tested whether students assigned "prek only group" and "follow-through group" were significantly different from each other, and control group children, on each of the listed characteristics.

Table 2. Head Start Impact Study Results

2a. Sustained classroom environment - End of Kindergarten Literacy Composite

	(1) End of HS	(2) Spring of K	(3) Spring of K	(4) Spring of K
Treatment	0.16*	-0.12+	-0.12+	-0.12+
	(0.07)	(0.06)	(0.07)	(0.07)
Total advanced literacy activities in K (times per month; standardized)			0.12*	0.10
			(0.05)	(0.07)
Total basic literacy activities in K (times per month; standardized)			-0.12*	-0.10
			(0.05)	(0.08)
Treat * Advanced literacy activities				0.03
				(0.09)
Treat * Basic literacy activities				-0.03
				(0.10)
Observations	1632	1077	1075	1075

Model 1 includes all children in the HSIS age-4 cohort; Models 2-4 only include children whose kindergarten teacher responded to the survey. Analyses of outcomes based on teacher response status available in Appendix C. Standard errors clustered at school level (in parentheses). All models are weighted using inverse probability of treatment weights to adjust for differential attrition and complex sampling. Weights include all the baseline child and family control variables. + p<.10; * p<.05; ** p<.01

2b. Sustained home environment -End of Kindergarten Literacy Composite

	(1) End of HS	(2) Spring of K	(3) Spring of K	(4) Spring of K	(5) Spring of K
Treatment	0.16*	-0.06	-0.10	-0.06	-0.11
	(0.07)	(0.06)	(0.07)	(0.09)	(0.33)
> High School deg.			0.43**	0.52**	0.50**
			(0.07)	(0.11)	(0.13)
Home literacy activities			0.15**	0.20**	
			(0.04)	(0.06)	
Home learning activities					0.02
					(0.02)
Treat * > High School deg.				-0.15	-0.16
				(0.15)	(0.16)
Treat * Home literacy activities				-0.09	
				(0.07)	
Treat * Home learning activities					0.01
					(0.02)
Observations	1632	1449	1449	1449	1449

Standard errors clustered at center of random assignment (in parentheses). All models are weighted using inverse probability of treatment weights to adjust for differential attrition and complex sampling. Weights include all the baseline child and family control variables. Home learning activities index and home literacy activities index were measured at the end of the HS year (standardized); items listed in Appendix A & B. + p<.10; * p<.05; ** p<.01

2c. Sustained classroom environment - End of First grade Literacy Composite

	(1) End of HS	(2) Spring of 1st Grade	(3) Spring of 1st Grade	(4) Spring of 1st Grade
Treatment	0.16*	-0.08	-0.07	-0.07
	(0.07)	(0.06)	(0.06)	(0.07)
Total advanced literacy activities in K and G1 combined (times per month; standardized)			0.12	0.12
			(0.07)	(0.11)
Total basic literacy activities in K and G1 combined (times per month; standardized)			-0.12+	-0.11
			(0.07)	(0.10)
Treat * Advanced literacy activities				-0.00
				(0.14)
Treat * Basic literacy activities				-0.02
				(0.12)
Observations	1632	1065	1065	1065

Model 1 includes all children in the HSIS age-4 cohort; Models 2-4 only include children whose first grade teacher responded to the survey. Analyses of outcomes based on teacher response status available in Appendix C. Standard errors clustered at school level (in parentheses). All models are weighted using inverse probability of treatment weights to adjust for differential attrition and complex sampling. Weights include all the baseline child and family control variables. + p<.10; * p<.05; ** p<.01

Table 3. Building Blocks Scale Up Study Results

3a. Sustained classroom environment – End of Kindergarten Math Composite

	(1) End of BB	(2) Spring of K	(3) Spring of K	(4) Spring of K	(5) End of BB	(6) Spring of K
Treatment	0.66***	0.33***	0.32***	0.33**	0.67***	0.32***
	(0.07)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)
Mathematics Teaching Quality			0.04	0.03		
			(0.04)	(0.06)		
Number of Math Activities			0.13*	0.13+		
			(0.05)	(0.07)		
Treat * Mathematics Teaching Quality				0.05		
				(0.08)		
Treat * Number of Math Activities				-0.02		
				(0.08)		
Treatment with Follow-Through					0.64***	0.38***
					(0.09)	(0.10)
Observations	555	555	555	555	865	865

Standard errors clustered at school level (in parentheses). Mathematics teaching quality and number of math activities were measured using the COEMET. For each variable, scores were averaged from classroom observations in the kindergarten year. All models include controls for gender, ethnicity, age at preschool entry, mother's education level, free or reduced price lunch status, special education status at preschool entry, whether limited English proficient, and blocking assignment. The sample was restricted to students non-missing on all preschool, kindergarten, and first grade mathematics measures, as well as the classroom observational measure in kindergarten and first grade (COEMET). + p<.10; * p<.05; ** p<.01

3b. Sustained home environment – End of Kindergarten Math Composite

	(1) End of BB	(2) Spring of K	(3) Spring of K	(4) Spring of K
Treatment	0.64***	0.32***	0.32***	0.34***
	(0.08)	(0.08)	(0.08)	(0.09)
> High School deg.			0.12+	0.14
			(0.06)	(0.12)
Home learning activities			0.04	0.01
			(0.05)	(0.08)
Treat * > High School deg.				-0.03
				(0.13)
Treat * Home learning activities				0.04
				(0.05)
Observations	555	555	555	555

Standard errors clustered at school level (in parentheses). Home learning activities was measured during the preschool year. All models include controls for gender, ethnicity, age at preschool entry, mother's education level (except for models 1 and 2), free or reduced price lunch status, special education status at preschool entry, whether limited English proficient, and blocking assignment. The sample was restricted to students non-missing on all preschool, and kindergarten mathematics measures, as well as the parent home survey. + p<.10; * p<.05; ** p<.01

3c. Sustained classroom environment – End of First Grade Math Composite

	(1) End of BB	(2) Spring of 1st Grade	(3) Spring of 1st Grade	(4) Spring of 1st Grade	(5) Spring of 1st Grade
Treatment	0.66***	0.15+	0.11+	0.11	0.17*
	(0.07)	(0.08)	(0.06)	(0.09)	(0.08)
Mathematics Teaching Quality			0.03	0.02	
			(0.04)	(0.06)	
Number of Math Activities			0.15**	0.21*	
			(0.05)	(0.08)	
Treat * Mathematics Teaching Quality				0.04	
				(0.09)	
Treat * Number of Math Activities				-0.09	
				(0.09)	
Treatment with Follow-Through					0.32***
					(0.09)
Observations		555	555	555	555	865

Standard errors clustered at school level (in parentheses). Mathematics teaching quality and number of math activities were measured using the COEMET. For each variable, scores were averaged from classroom observations in the kindergarten and first grade year. All models include controls for gender, ethnicity, age at preschool entry, mother's education level, free or reduced price lunch status, special education status at preschool entry, whether limited English proficient, and blocking assignment. The sample was restricted to students non-missing on all preschool, kindergarten, and first grade mathematics measures, as well as the classroom observational measure in kindergarten and first grade (COEMET). + $p<.10$; * $p<.05$; ** $p<.01$

Appendix A. Coding scheme for instructional quality of literacy activities in the Head Start Impact Study

Kindergarten literacy activities		First grade literacy activities	
Listen to stories with no print	basic	Activity related to book	basic
Show child how to read a book	basic	Write letters of alphabet	basic
Write own name	basic	Learn names of letters	basic
Teach directional words like over and up	basic	Have children tell you a story	basic
Write letters of the alphabet	basic	Practice sounds letters make	basic
Learn the names of letters	basic	Listen to stories w. print	basic
		Read books chosen by child	basic
		Read text w controlled vocab	basic
		Read text w strong phonemic pattern	basic
		Read patterned or predictable text	basic
		Hear storytellers	basic
Discuss new words	advanced	Language activities in mixed achievement groups	advanced
Have children tell you a story	advanced	Discuss new words	advanced
Practice the sounds that letters make	advanced	Read aloud	advanced
Listen to stories with print	advanced	Read silently	advanced
Rhyming words and families	advanced	Work in reading workbook	advanced
		Write words from dictation	advanced
		Use invented spellings	advanced
		Read thematic text	advanced
		Compose stories or reports	advanced
		Publish child's writing	advanced
		Perform plays/skits	advanced
		Write stories in journal	advanced

Appendix B: Items for HSIS Parent Survey included in the Home Literacy and Home Learning Activities Indices

Home Literacy Activities Index	Home Learning Activities Index
Items coded on 1-4 likert scale	*Items coded 1=true or present in home; 0= not*
Number of times child is read to	Follow a daily routine
Work on learning names of letters	Keep notes about behavior
Practice letters of the alphabet	Collect sample of child's work
Discuss new words	Collecting photos
Have [child] tell you a story	Chart behavior/skills
Practice sound of letters	Gone to a movie
Listen to stories with print	Gone to a play or concert
Listen to stories not seeing print	Visited art gallery or museum
Retell or make up stories	Visited playground or park
Show child how to read	Talked with child about heritage
Child practices writing/spelling name	Attend community events
Learn about rhyming words	Takes child along on errands
How often you read books	Comic books in home
	Books for children in home
	Magazines for children in home
	Magazines for adults in home
	Newspapers in home
	Catalogs in home
	Religious books in home
	Dictionaries or encyclopedias in home
	Other books in home
	Take books home from library

Appendix C: Head Start Impact Study- Fadeout estimates by Kindergarten and First Grade teacher survey response

	(1) End of HS	(2) End of HS; K Teacher survey missing	(3) End of HS; K Teacher survey nonmissing	(4) Spring of K	(5) Spring of K; K Teacher survey missing	(6) Spring of K; K Teacher survey nonmissing	(7) Spring of 1st Grade	(8) Spring of 1st Grade; G1 Teacher survey missing	(9) Spring of 1st Grade; G1 Teacher survey nonmissing
Treatment	0.16*	0.16	0.17+	-0.05	0.06	-0.12+	-0.05	-0.10	-0.08
	(0.07)	(0.14)	(0.09)	(0.06)	(0.17)	(0.06)	(0.06)	(0.24)	(0.06)
Observations	1632	589	1043	1525	448	1077	1313	248	1065

Standard errors clustered at center of random assignment (in parentheses). All models are weighted using inverse probability of treatment weights to adjust for differential attrition and complex sampling. Weights include all the baseline child and family control variables. + p<.10; * p<.05; ** p<.01

Appendix D. Additional models testing sustained classroom environment in the Head Start Impact Study: Kindergarten Literacy Composite Scores

	(1) Kindergarten classroom fixed effect	(2) Full-Day Kindergarten	(3) Kindergarten class size	(4) Classroom-level poverty (FRPL)
Treat	0.02	-0.23*	-0.38	-0.10
	(0.32)	(0.11)	(0.32)	(0.20)
Low literacy composite score (std.)				
High literacy composite score (std.)				
Treat * high literacy composite score at K entry				
Treat * low literacy composite score at K entry				
Full-day K		0.23		
		(0.17)		
Treat * Full-day K		0.19		
		(0.14)		
Class size			0.01	
			(0.01)	
Treat * Class size			0.01	
			(0.01)	
Classroom % FRPL				0.02
				(0.27)
Treat * Class % FRPL				-0.34
				(0.28)
Observations	1077	1008	971	777

Standard errors clustered at school level (in parentheses). All models include fixed effects for center of random assignment and are weighted using inverse probability of treatment weights to adjust for differential attrition and complex sampling. Weights include all the baseline child and family control variables. Changes in observation counts across models reflect changes in teacher survey item non-response. FRPL-Free and Reduced-price lunch. + p<.10; * p<.05; ** p<.01

Appendix E. Head Start Impact Study: Selection into classroom and school environments - Kindergarten

	(1) Total advanced literacy activities (times per month; standardized)	(2) Total basic literacy activities (times per month; standardized)	(3) Yrs. teaching exp.	(4) HS or below (Teacher)	(5) Some college (Teacher)	(6) Associates (Teacher)	(7) College (Teacher)	(8) College+ (Teacher)	(9) Pre-k teaching license	(10) Elementary teaching license	(11) Full-day K
Treatment	-0.029	-0.100	-0.803	0.002	0.010	0.001	-0.012	-0.000	-0.025	0.001	-0.052**
	(0.07)	(0.07)	(0.74)	(0.00)	(0.01)	(0.01)	(0.03)	(0.03)	(0.04)	(0.01)	(0.02)
Observations	1075	1075	1062	1071	1071	1071	1071	1071	1003	1003	1008

	(12) Classroom num LEP students	(13) Classroom num FRPL eligible	(14) Class size	(15) Teaching assistant	(16) Percent of school children black	(17) Percent of school children eligible for free/reduced lunch	(18) Percent of school children Hispanic	(19) Percent of school children white	(20) School proficiency level in math	(21) School proficiency level in reading
Treatment	0.041	-0.188	-0.325	-0.010	0.005	0.018	0.032	-0.033	-1.608	-1.309
	(0.50)	(0.45)	(0.34)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(1.55)	(1.85)
Observations	1006	821	971	992	925	898	925	925	928	927

Standard errors clustered at center of random assignment (in parentheses). All models are weighted using inverse probability of treatment weights to adjust for differential attrition and complex sampling and include fixed effects for center of random assignment. Weights include all the baseline child and family control variables. + p<.10; * p<.05; ** p<.01